

Teasing apart sources of stochastic variations in eukaryotic gene expression

Teresa Przytycka

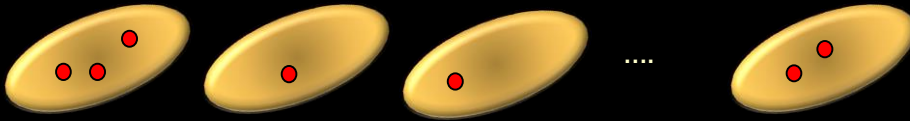
AlgoCSB

***Algorithmic Methods in
Computational and Systems
Biology***

NCBI / NLM / NIH



Stochastic variability of gene expression in isogenic cell population “expression noise”



- isogenic cells
- single cell measurements

Noise measure

δ^2 - variance
 μ - mean

Noise a phenotype that can be regulated

Noise can be beneficial – increase evolutionary plasticity

Noise might be harmful for a response that needs to be tightly regulated

Sources of gene expression noise

Intrinsic noise – gene specific related to stochasticity of the gene expression process
(transcription, translation, degradation of mRNA and protein)

Extrinsic noise – environment, cell size, noise propagated through interaction network ...

Random Poisson process

Noise scaling with abundance

$$CV^2 = 1/\mu$$

Fano Factor (Fano Noise)

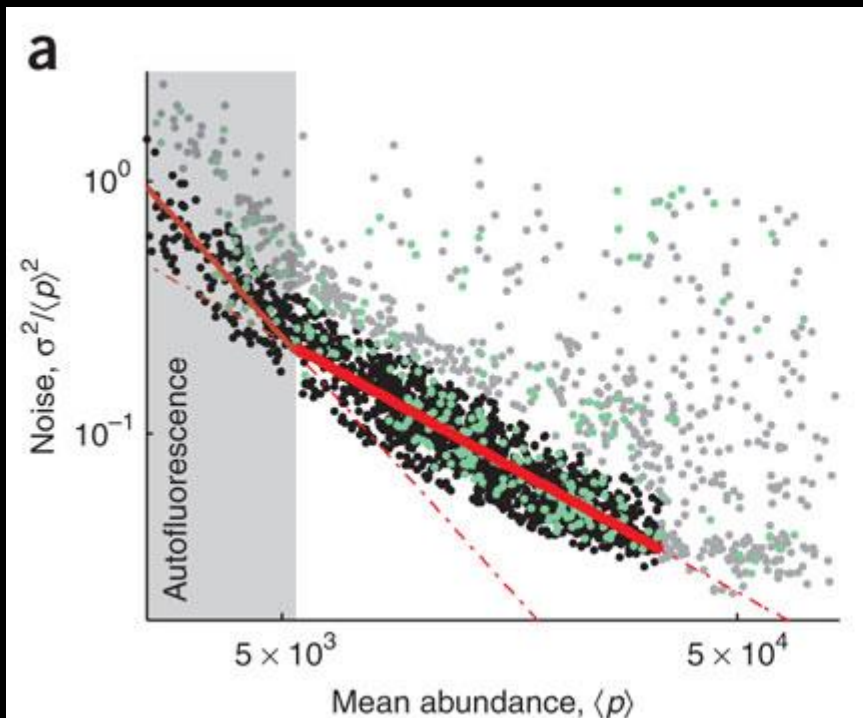
$$F = CV^2 \mu$$

For Poisson process:

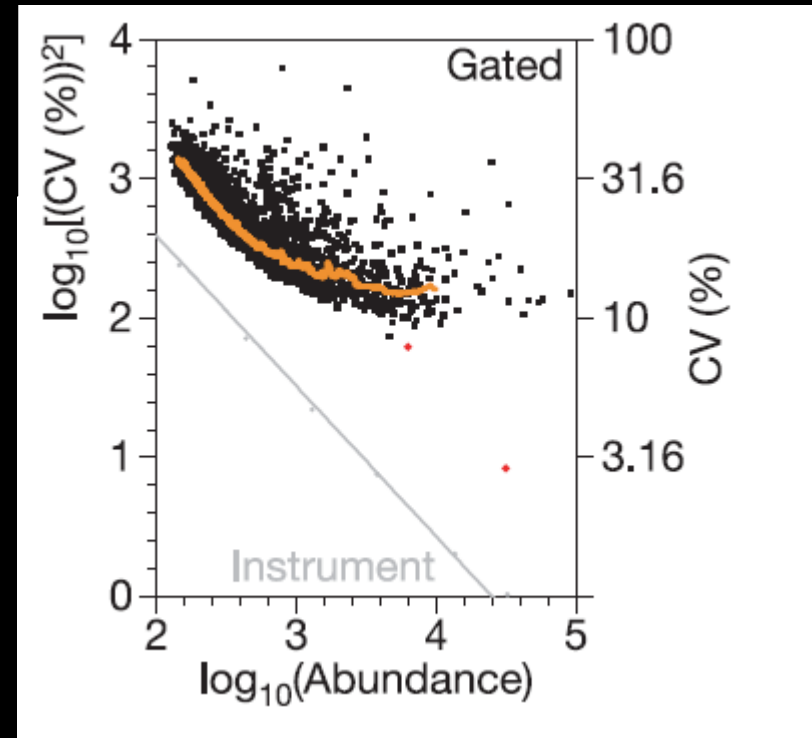
$$F = 1$$

As expected from a random stochastic process noise (squared) scales with abundance

Bar-Even, A. et al. Noise in protein expression scales with natural protein abundance. *Nat Genet*, 2006



Newman, J.R. et al. Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature* 441, 2006.



- Gene specific noise properties: **divergence from trend line**

Measuring divergence from trend line

- DM measure (Newman et al.)
 - Euclidean distance (in y-direction)
- Noise differential (Salari et al.)
 - ratio of Fano factors
- Noise residua (Bar-Even et al.)
 - \log (noise differential)

Expectations based on theoretical models

Findings on Newman *et al.*

- Efficient transcription and transcription bursts should correlate with increased noise

highly statically significant correlation
between TATA box and increased noise

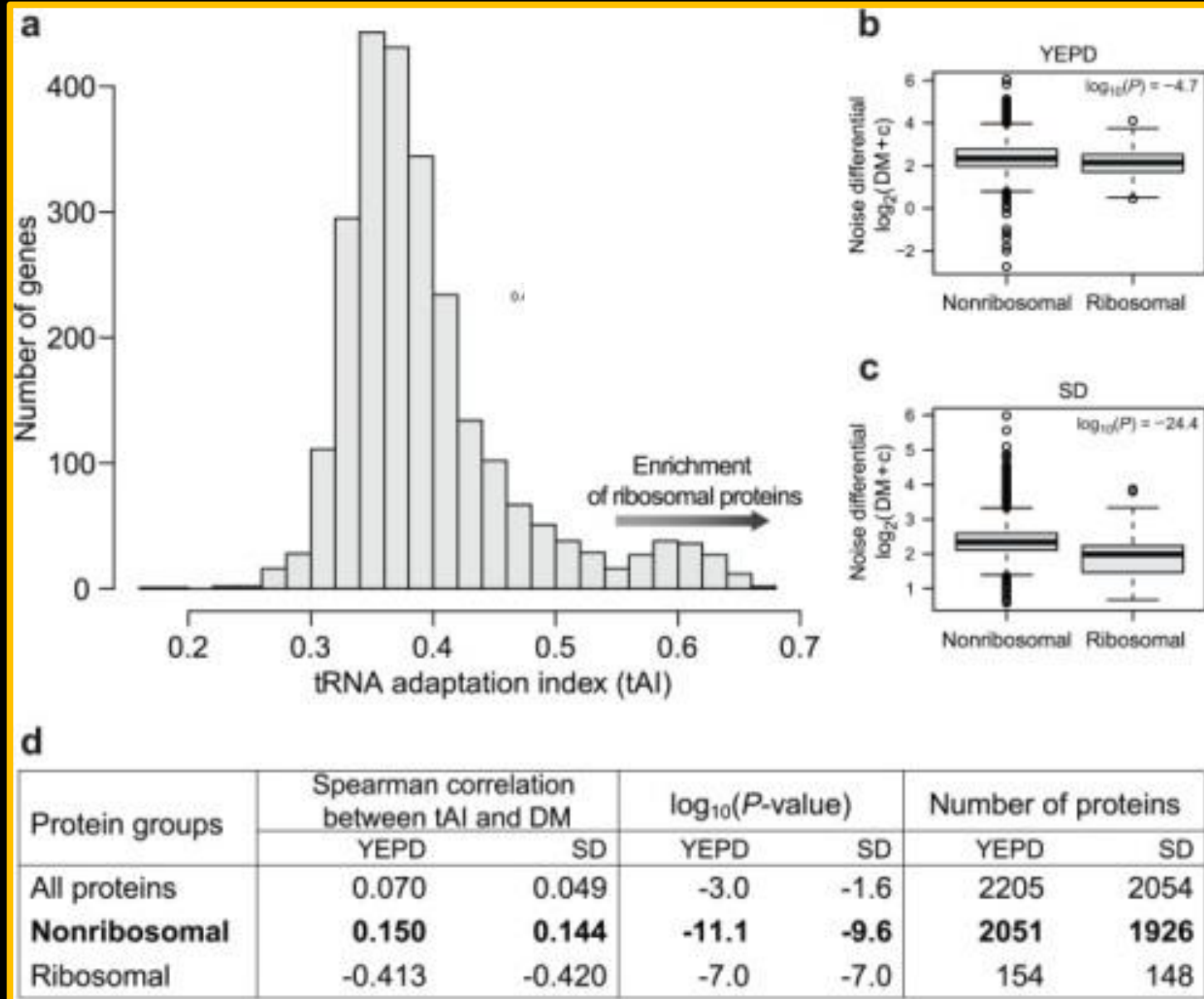
- Efficient translation (multiple proteins form the same transcript) should correlate with increased noise

No highly statically significant correlation
between Codon Usage (tAI – tRNA Adaptation Index) and noise

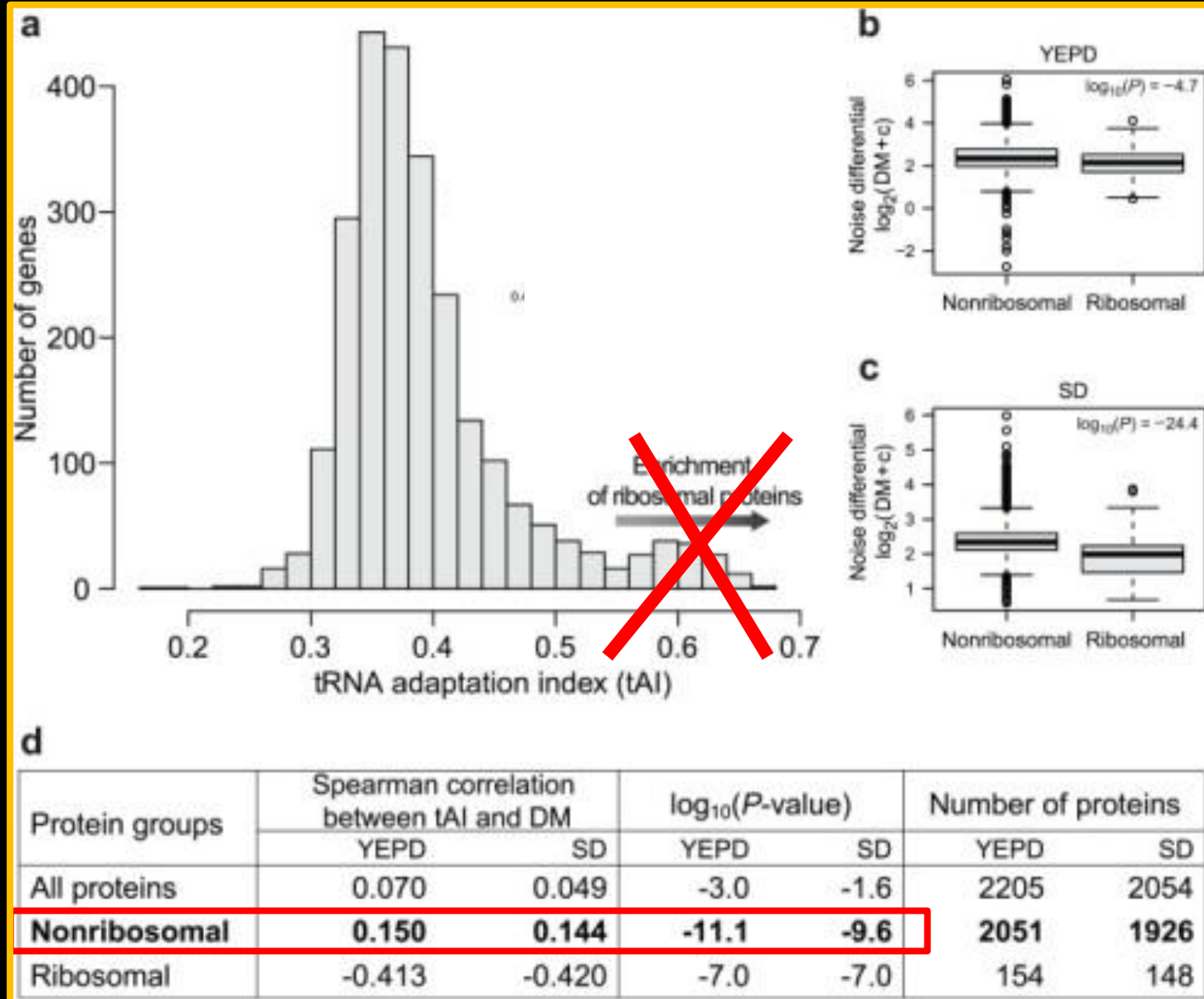
Conjecture:

This lack of correlation is due to averaging effect of various stochastic processes across all genes while the contribution of these processes might be different in different groups

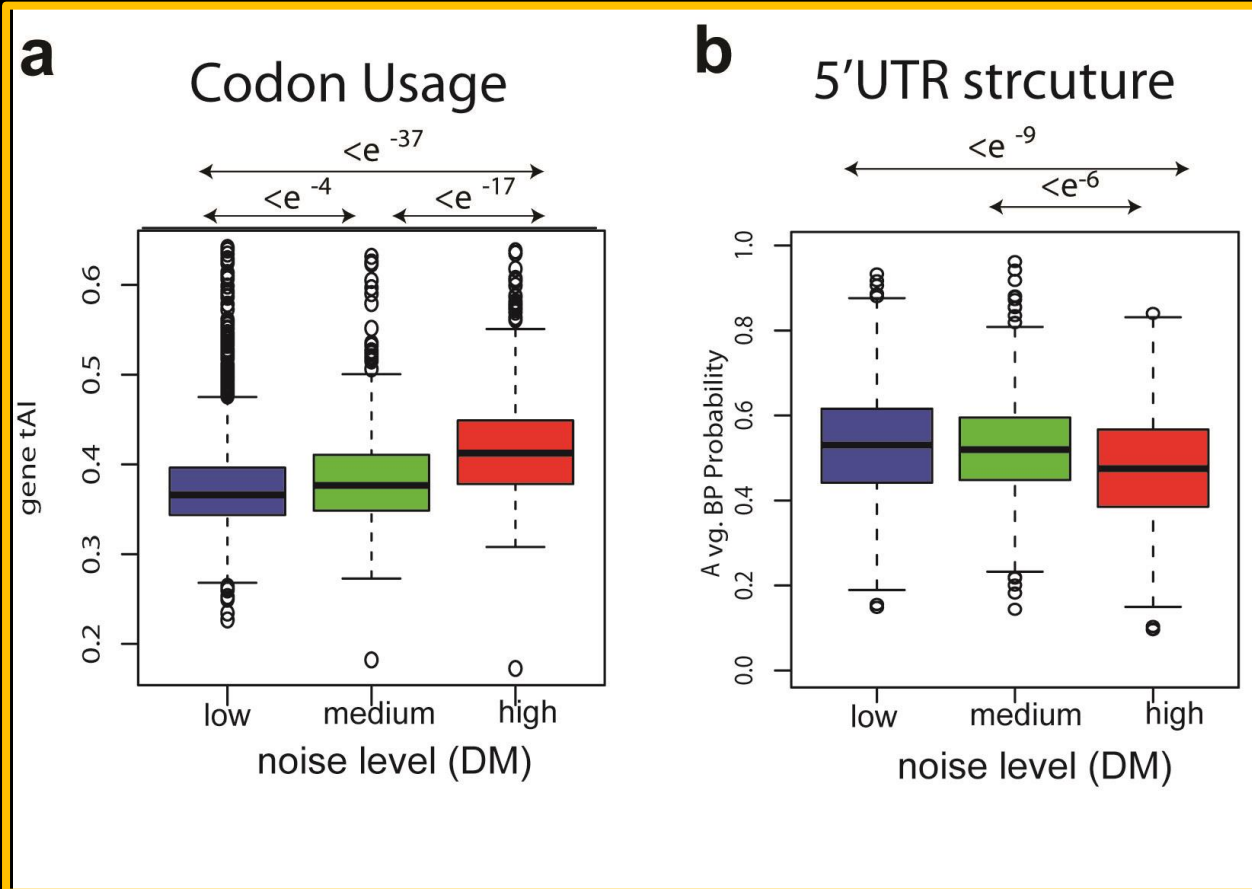
Two populations of genes based codon adaptation index



Two populations of genes based codon usage



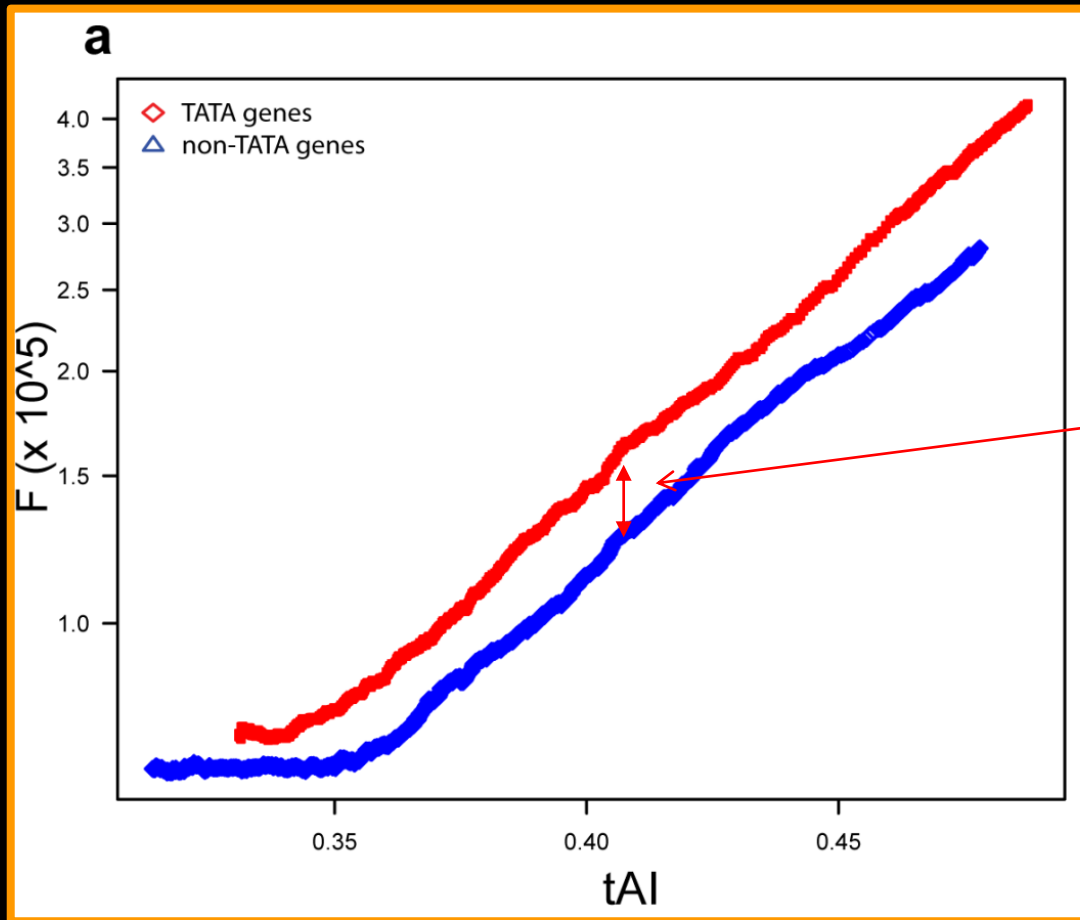
Noise level (DM) is correlated with translation related features



Looking at the interplay between transcription and translation related features

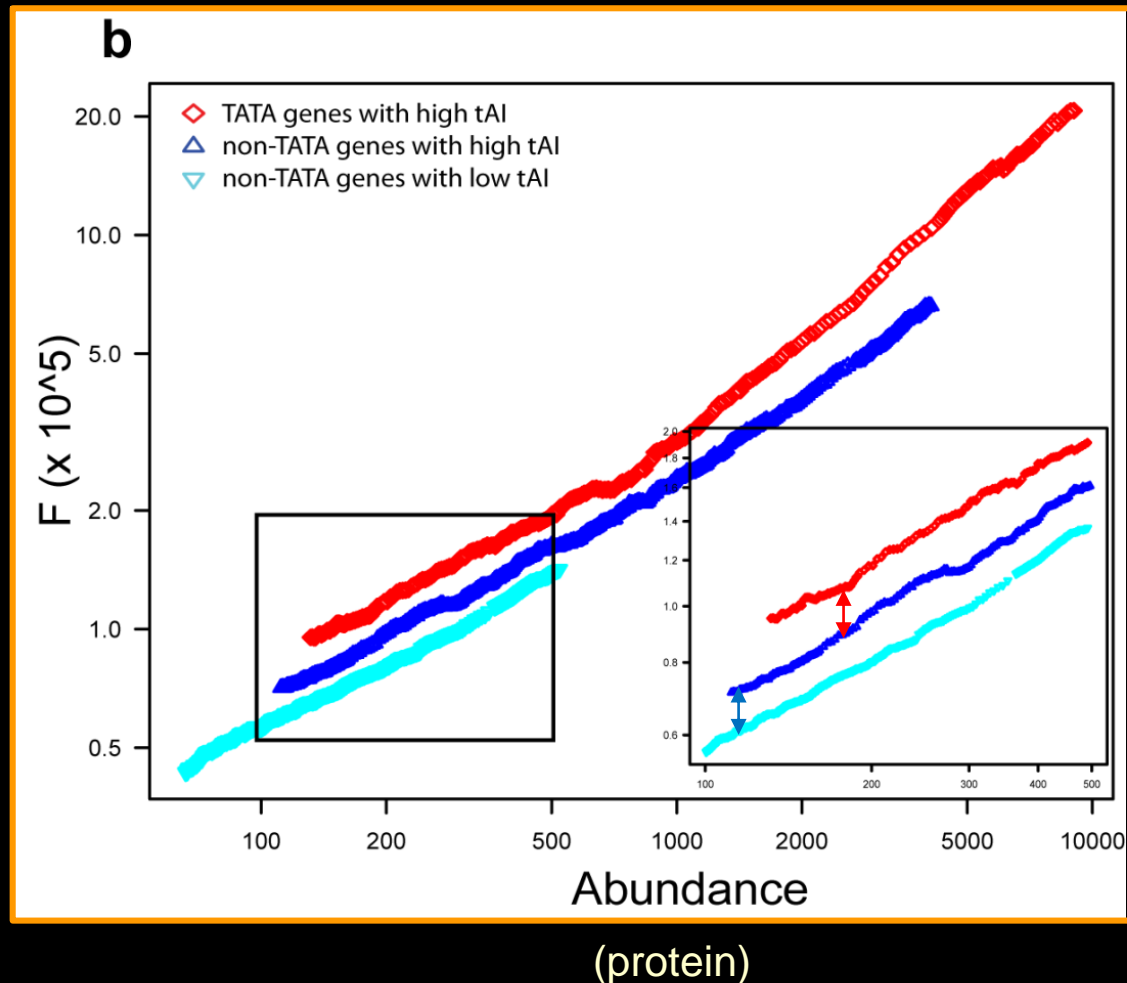
- Transcription – TATA box
- Translation – codon usage (tAI measure)

Compression of the relation between Fano factor and codon usage for TATA and no-TATA genes

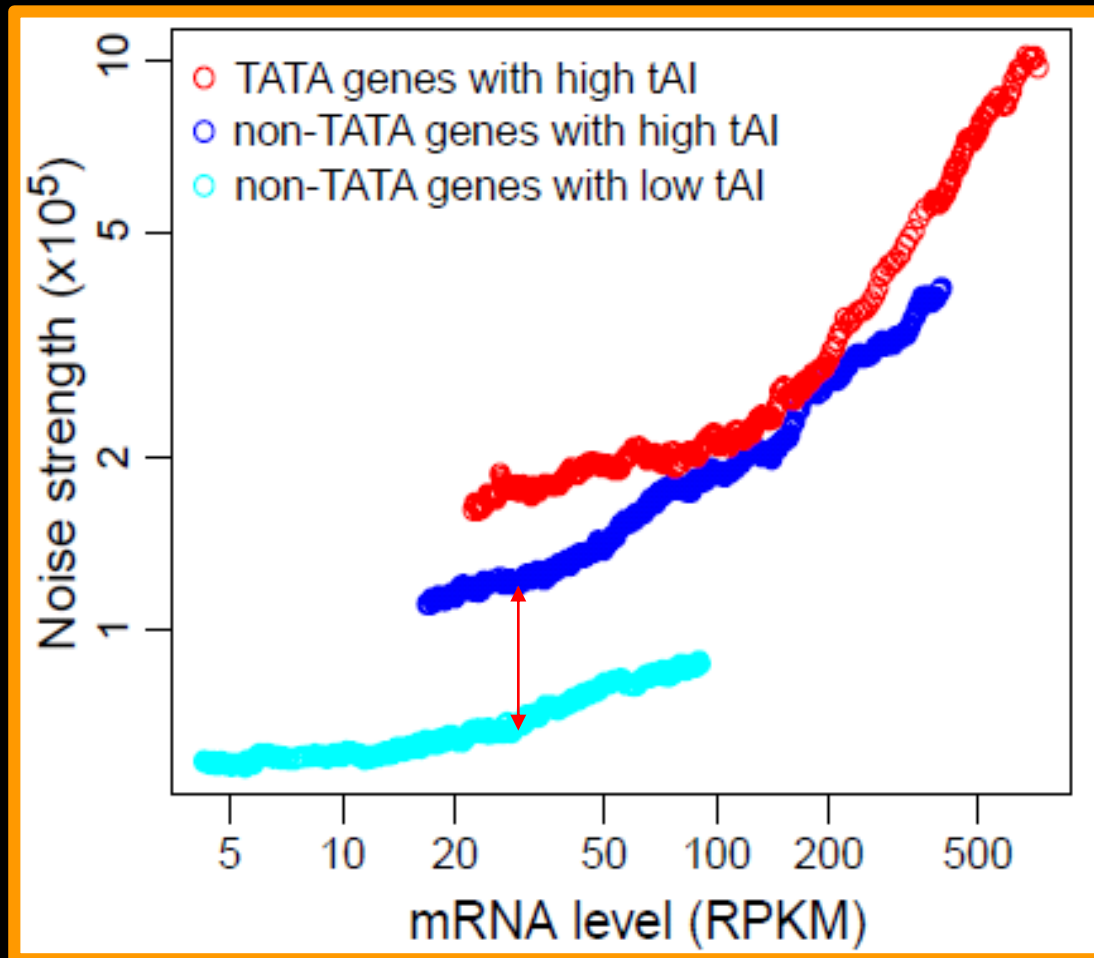


Impact of TATA box

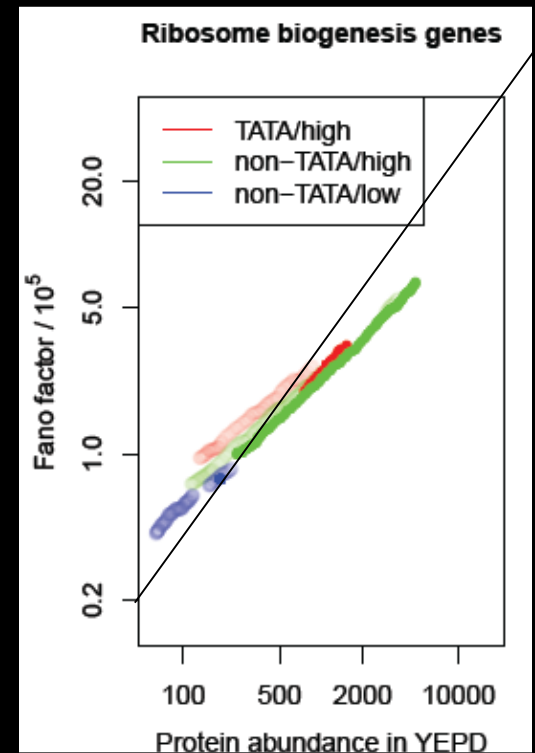
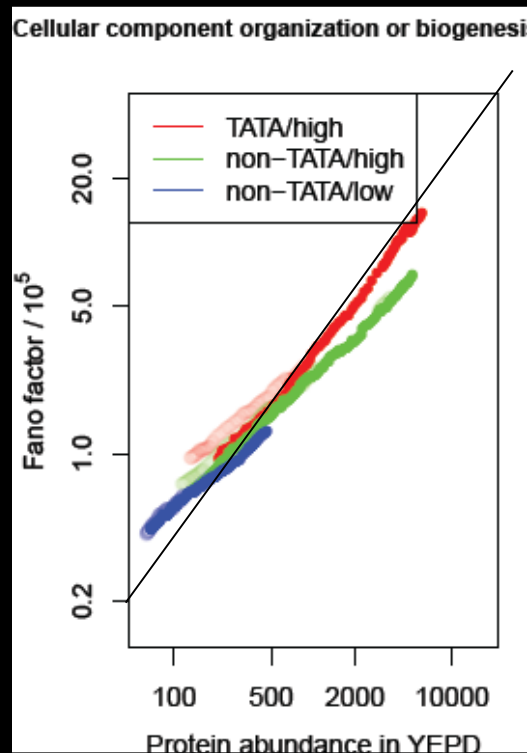
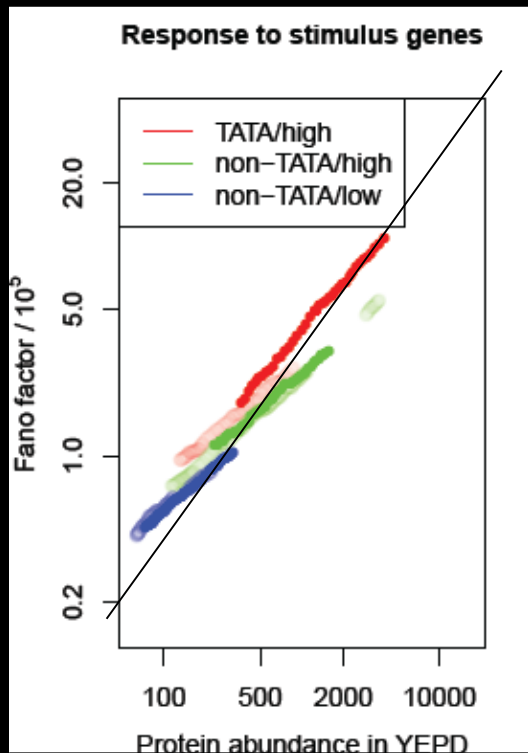
The impact of translation efficiency (measured as noise differential) is of the same magnitude as of transcription efficiency



High/low tAI gap is more pronounced when normalized by mRNA level



Noise regulation within large GO groups



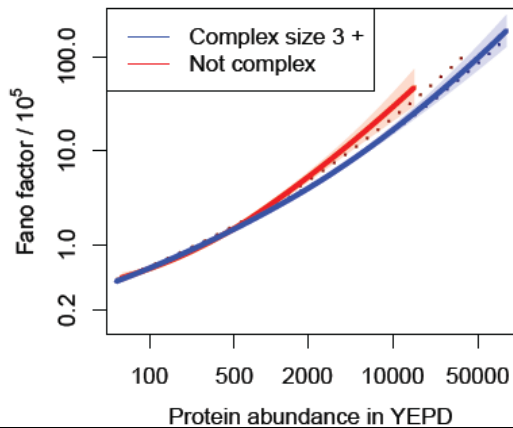
high noise GO group

low noise GO group

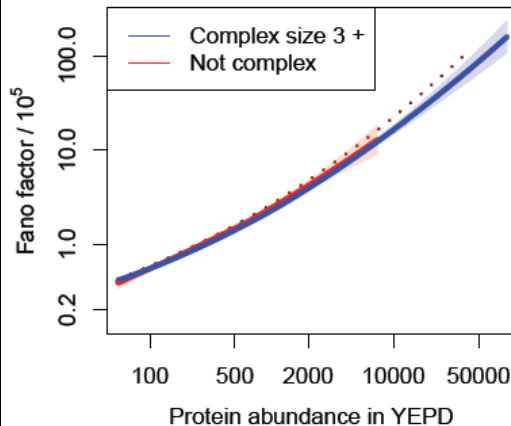
Examining the hypothesis that genes participating in complexes are less noisy

- In Newman *et al.* studies: no genome-wide correlation between complex participation and noise
- Our studies: indication of such relation within individual functional groups

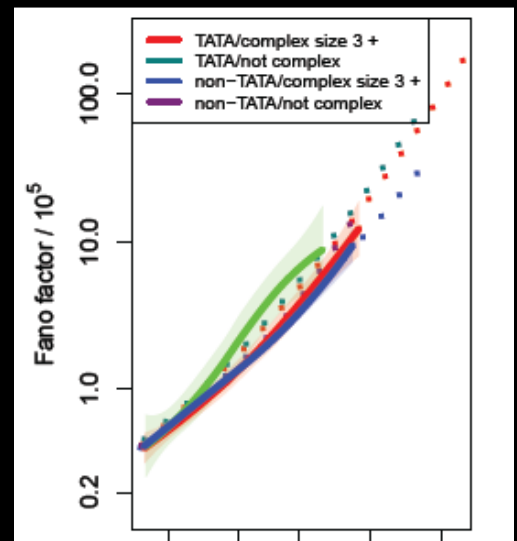
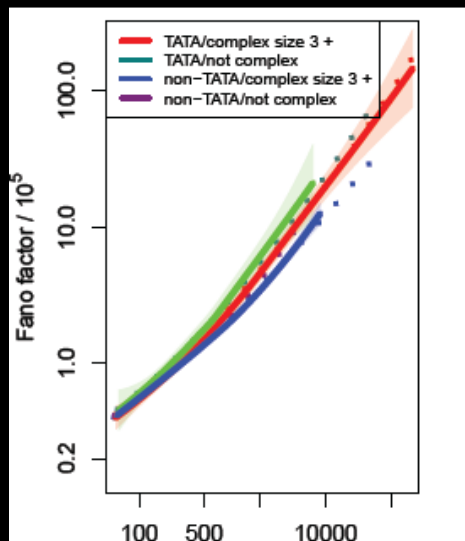
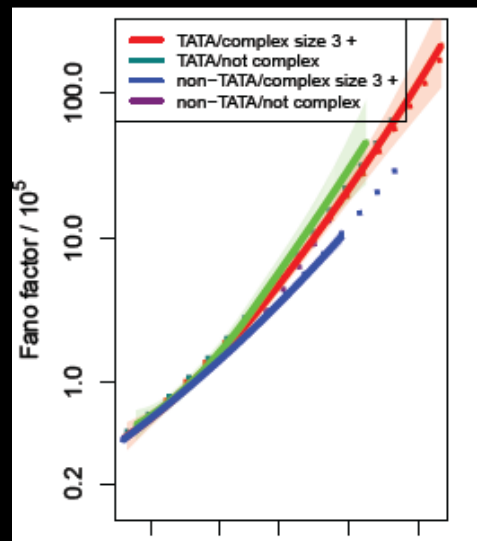
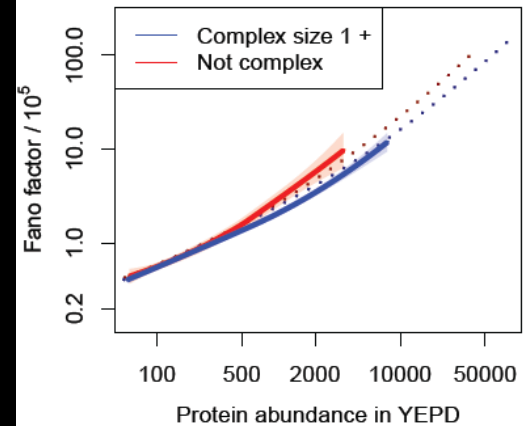
Cellular component organization or biogenesis genes



Regulation of biological process genes



Negative regulation of biological process genes



Summary

- Genomic features associated with efficient translation associate with increased noise
- Noise level is defined by a complex trade-off which cannot be completely captured looking at genome-wide average behavior and it is often helpful to look at it from the perspective of individual functional gene groups

Acknowledgments

- Raheleh Salari, AlgoCSB/NCBI currently Stanford U.
- Jie Zhang, AlgoCSB/NCBI currently Nanyang TU, Singapore
- Damian Wojtowicz, AlgoCSB/NCBI, NIH
- David Levens, NCI, NIH
- Yitzhak Pilpel, Weizmann Inst. Rehovot, Israel
- Daniela Ganelin, AlgoCSB/NCBI and Blair High School